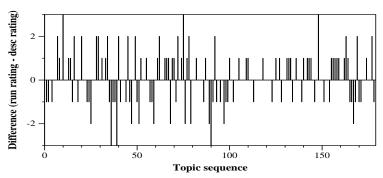
Summary Statistics			
Run ID	categoryaware1		
Task	Summarization		
Run type	automatic		
Summarization model	Supervised neural or deep learning model; Transfer learn-		
	ing or other pre-trained model		
Summarization method	Abstractive		
Number of summaries	179		

Overall measures	
Summaries rated "excellent"	26 (14.5%)
Summaries rated "good"	49 (27.4%) 69 (38.5%) 35 (19.6%)
Summaries rated "fair"	69 (38.5%)
Summaries rated "bad"	35 (19.6%)

ROUGE score compared to episode description		
	0.2270 (95%-conf.int. 0.21600 - 0.23945)	
ROUGE_L R score	0.2082 (95%-conf.int. 0.19889 - 0.21721)	
ROUGE_L F score	0.1762 (95%-conf.int. 0.16792 - 0.18456)	

Qualitative judgments		
Q1: Does the summary include names of the main people	64.2%	
(hosts, guests, characters) involved or mentioned in the		
podcast?		
Q2: Does the summary give any additional information	40.2%	
about the people mentioned (such as their job titles, bi-		
ographies, personal background, etc)?		
Q3: Does the summary include the main topic(s) of the	76.5%	
podcast?		
Q4: Does the summary tell you anything about the format	59.2%	
of the podcast; e.g. whether it's an interview, whether it's		
a chat between friends, a monologue, etc?		
Q5: Does the summary give you mode context on the title		
of the podcast?		
Q6: Does the summary contain redundant information?		
Q7: Is the summary written in good English?		
Q8: Are the start and end of the summary good sentence		
and paragraph start and end points?		



Per-topic difference in rating from episode description